T
HZ:538
.2

SCHS STATISTICAL PRIMER

State Center For Health Statistics

Department of Human Resources • Division of Health Services • P.O. Box 2091 • Raleigh, NC 27602-2091 • 919/733-4728

N. C. DOCUMENTS

DEC 3 1986

N. C. STATE LIBRARY
RALEIGH

No. 7 October 1986

QUESTIONNAIRE DESIGN

by

J. Michael Bowling

ABSTRACT

An overview of the survey process is presented with special emphasis on questionnaire design. Topics covered are types of questions, structure of questions, question wording, and questionnaire construction. This information should help persons in state and local agencies to collect better data through more effective questionnaire design and implementation.

The collection, processing, and dissemination of information have become increasingly important to the health care provider, researcher, and administrator faced with competing alternatives for the expenditure of scarce resources. Sample survey design, once relegated by many students to the dusty shelf of pedagogic esoterica, has become a useful tool to a more diverse group of data users employing more sophisticated computer technologies. The intent of this *Statistical Primer* is to provide this newly emerging audience of data users with an overview of the "survey process" placing special emphasis upon questionnaire design.

Conducting A Survey

A survey is a systematic means of collecting information directly from people (respondents) often through the use of a questionnaire. A census is a special type of survey in which efforts are made to gather information from all members of the population. In contrast, a sample survey normally involves the selection of a small percentage of eligible members of a population representative of the larger group. If appropriate steps have been taken to eliminate bias in the selection of sample respondents, the sample size and variability of the sampling units can be used to determine the precision of estimated means, proportions, and other population indices being measured in the survey.

There are three basic types of surveys: face-to-face, telephone, and mail. The choice of method is dependent upon, among many other things, the research goal of the survey, the population chosen for study, and the survey budget.

Warwick and Lininger (1) have delineated seven interdependent stages in conducting a sample survey, most of which also apply in the case of a census of all eligible respondents. The following is an overview of these stages.

Planning is the preliminary stage of the survey. During the planning stage researchers should conduct an extensive literature review of previous publications to be used as a basis for questionnaire design. The formulation of a "research question" as a focus for the project is often helpful in ensuring continuity between the stages of the survey which culminate in data analysis and report writing.

Sampling is the process of selecting a subset of eligible respondents from a population. The key to the cost-effectiveness of choosing a small number of respondents and generalizing to a much larger whole is the representativeness of the sample. Consultation with a survey statistician, while beneficial to all stages of survey design, is most crucial in the sampling stage. The foundation of a good sample survey rests upon the sample (1).

Questionnaire design is the stage of the survey process in which the goal of the project, often embodied in the research question, is translated into a series of questions. The principles of designing a survey questionnaire are the same whether the scope of a survey project is limited to an interoffice canvas or whether it is a multistage national survey. A discussion of these principles is presented in most of the remainder of this report.

Data collection includes the recruitment and training of interviewers as well as the actual collection of data. Interviewer training and data collection may be the most neglected and often least organized of the efforts in the survey project. Well-trained interviewers, using methods designed to enhance respondent interest in the survey, are more likely to obtain accurate responses from sample members and will often operate more efficiently in completing their work. The success of mail surveys is particularly dependent upon the "packaging" of the questionnaire in a pleasing and interesting manner.

Editing and coding of the questionnaire are necessary steps in ensuring that answers have been properly entered on the questionnaire form and are converted to codes that can be entered into a computer. Only on those rare occasions in which the size of the study is very limited (less than 25 respondents) and the need for subgroup analysis is minimal would the time and expense of computer processing outweigh the advantages of computer tabulation over hand tabulation.

Preparation for analysis includes a variety of steps like keypunching data, performing range and consistency checks on the responses, and weighting the data to reflect the method of sampling and limitations of the survey (e.g., nonresponse and incomplete sample coverage).

Analysis and report writing vary considerably by the sophistication of the researcher and the audience, as well as by the demands of the research question. For example, simple graphs or descriptive charts are often sufficient to confirm expected relationships, while complex regression analysis may be needed to establish those which are less obvious. Successfully meeting analysis goals always requires that the research question of the study be identified during the planning stage.

Questionnaire Design

The successful completion of a survey is the culmination of efforts in each of the seven stages listed above. The science of survey sampling has been advanced to the point that appropriate sampling strategies have been identified for most survey settings. Procedures for data collection, coding, editing, and analysis, although still the object of investigation by survey methodologists, are also rather well established.

Questionnaire design, the remaining stage, is the area that may be among the most amenable to improvement in many survey designs. Thus the remainder of this report will present guidelines for the construction of a survey questionnaire.

Writing questions is probably one of the least precise aspects of survey design. Whereas many books have been published that provide a "cookbook" approach to other aspects of the survey process, question writing, at best, is dealt with by vague prescriptions and recommendations of structure. Given the variety of subject matter encountered in survey research, researcher discretion remains the most important input to questionnaire design. This section will highlight some problematic areas of question writing and suggest alternatives for overcoming them.

Dillman (2) suggests that the writing of questions can be divided into three decisions: 1) What type of information is sought? 2) What question structure is most effective in gathering the information? 3) What is the choice of words for the question?

Types of Questions

Initially, each researcher must consider what type of information is needed from a particular question. Most questions address the following information types: attitudes, beliefs, behavior, and attributes.

Attitudes are the dispositions, either favorable, unfavorable, or neutral, that individuals hold towards "attitude objects". These objects may be highly charged concepts such as abortions, or less intense concepts such as the election of state judges.

Example 1: Questions Designed to Elicit Attitudes

Q1.1 In general how do you feel about the Equal Rights Amendment? Do you strongly favor, somewhat favor, somewhat oppose, or strongly oppose the amendment?

1. Strongly favor 1
2. Somewhat favor 2
3. Somewhat oppose 3
4. Strongly oppose 4
5. No opinion 5

Q1.2 Should organized prayer be allowed in public schools?

1. Yes 1
2. No 2
3. Don't know . . . 3

Beliefs are facts that individuals hold as true or false. The feelings attached to the belief are ideally not elicited in a belief question. Notice the examples presented below. Knowledge that the ERA is supported by women's organizations such as NOW does not indicate the respondent's attitudes towards NOW or the ERA. Attitude questions are most appropriate for discovering the respondents feelings towards the two attitude objects.

Example 2: Questions Designed to Elicit Beliefs

Q2.1 Is the following statement true or false?

The woman's organization NOW supports the adoption of the Equal Rights Amendment.

1. True 1
2. False 2
3. Don't Know . . 3

Q2.2 How important is the abortion issue to you—would you say it is very important, important, not very important, or not important at all?

1. Most important 1
2. Important 2
3. Not very important 3
4. Not important at all 4

The third category of information elicited in survey questions is behavior. This category includes behaviors respondents have or have not performed. Behavioral questions may reference the past, present, or future.

Example 3: Questions Designed to Elicit Behavior

Q3.1 Are you currently trying to lose weight?

1. Yes 1
2. No 2

Q3.2 How often do you eat breakfast—almost every day, sometimes, rarely, or never?

1. Almost every day 1
2. Sometimes 2
3. Rarely 3
4. Never 4

Attributes are factual characteristics of respondents often collected in survey research. Common attributes of interest include sex, income, education, marital status, race, and religion. Attribute questions are normally placed at the end of a questionnaire, with income often the last question. More sensitive than most survey topics, attribute questions are more likely to elicit respondent anxiety and hesitance to continue. If this results in a partial interview, more data are saved with this placement strategy.

Example 4: Questions Designed to Elicit Attributes

Q4.1 Last week were you working full-time, part-time, going to school, keeping house, or what?

1. Working full-time 1
2. Working part-time 2
3. With a job, but not at work because of temporary illness, vacation, strike 3
4. Unemployed, laid off, looking for work . . . 4
5. Retired 5
6. In school 6
7. Keeping house 7
8. Other: SPECIFY _____ 8

Q4.2 What is your present age?

_____ AGE

Structure of Questions

After the survey researcher has decided what type of question is needed, the next decision to consider is the structure of the question. Question structure is based upon the degree of discretion allowed both respondents and interviewers in choosing answers to questions. The open-ended question allows complete freedom of expression in answering questions. Respondents are not prompted to answer in a predetermined manner and are often instructed to use their own words.

The open-ended question is a useful tool in gathering information that requires free association on the part of the respondent or involves a multitude of possible response categories. Exploratory studies may rely upon open-ended questions with the use of interviewer probes to gather detailed information. These data can then be used to develop a more structured questionnaire or form the content of a case-study design.

Some open-ended questions written for telephone or in-person surveys include possible response categories that are silently coded by the interviewer. The classification of possible responses, with a residual "other" category, tends to shorten the interview and improve comparability to responses across respondents. Example 5 provides an open-ended question with and without response classification.

Example 5: Open-Ended Questions

Q5.1 What kind of work do you normally do? That is, what is your job called?

_____ OCCUPATION

Q5.2 Now let's move on to some questions about health information. What would you say is your **most** important source of information about health care?

a. Doctor 01
b. Public Health Department 02
c. Other clinic or health center 03
d. Other people—family 04
e. Other people—friends, acquaintance 05
f. Newspapers 06
g. Television 07
h. Magazines 08
i. Radio 09
j. Other sources: SPECIFY _____ 10
k. None 00
l. DK/NA 99

Answers to open-ended questions without classification such as Q5.1 are recorded either verbatim or paraphrased by the interviewer, depending upon the nature of the question. Occupation is often asked in an open-ended question, with numeric coding conducted afterwards.

An open-ended question with classification is provided in Q5.2. The response categories allow interviewers to code volunteered responses without prompting respondents on possible sources of health information. In many cases respondents have never considered their sources of health information and some concentration is necessary to formulate a response. Respondents should never be prompted in such an open-ended question to aid them in formulating a response.

Close-ended questions provide the respondent with a list of possible answers from which to choose the best or in some cases multiple responses. The advantages of close-ended questions are that they are easier to answer, requiring much less concentration on the part of the respondent; easier to code, involving much less interviewer discretion in choosing the appropriate response category; and most importantly, the close-ended question provides comparable data across respondents.

In designing questionnaires, researchers should consider the demands being placed upon respondents. Open-ended questions requiring considerable concentration can tax respondents to the degree that they will either prematurely terminate the interview or hurriedly answer in order to complete the questionnaire. A mix of close-ended and open-ended questions can be used to limit demands on the respondent.

Ease in coding responses by the interviewer or respondent is an important consideration in questionnaire design. Pauses between questions, the result of lengthy recording of responses, increases the time required to complete the interview. Inexperienced interviewers are likely to encounter difficulty "fitting" responses to open-ended questions into available response categories.

Comparability across respondents is an important prerequisite to combining answers and calculating aggregate summary statistics. The use of open-ended questions with classification of responses does not fully eliminate the problem of comparability. Consider Q5.2 in which respondents are asked to recall sources of health information. To the degree that recollection, understanding, and articulation play a part in the respondent's formulation of an answer, comparability across individuals is hampered. Ideally, comparisons should be among respondents' actual source of information and not their recollection or understanding of the source.

There are two types of close-ended questions: ordered response and unordered response. Q6.1 is an example of an ordered-response question. In this type of question the possible responses fall in an order along a continuum of potential responses. The number of response categories in an ordered response question is dependent upon the goal of the question and the type of survey (i.e., mail, in-person, telephone). Mail and in-person surveys can use greater detail in the categories through the use of visual displays of the responses. Telephone surveys are limited by the respondent's memory of response options.

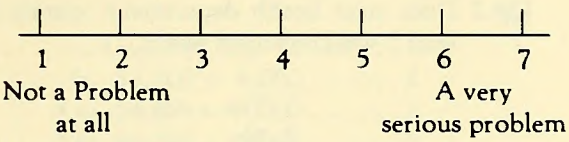
In Q6.2 an alternate display of a close-ended question with an ordered response is provided. Here respondents are allowed to choose from a numerical range. This form is limited mainly to in-person and mail surveys.

Example 6. Close-ended Questions

Q6.1 Would you say that the children in your household have gone boating often, seldom, or not at all in the past few years?

a. Often 1
b. Seldom 2
c. Not at all 3
d. Don't know 8

Q6.2 Below is a list of potential problems that teachers may have encountered in the fluoride mouthrinse program. Some teachers may consider the statement not a problem at all. Suppose these people are at one end of the scale at point number 1. Other teachers may consider the statement a very serious problem. Suppose these people are at the other end, at point 7. And of course, other teachers have opinions somewhere in between at point 2,3,4,5 or 6. Where would **you** place yourself on this scale for . . . . .?



A. The fluoride mouthrinse program takes too much time.

Q6.3 Think about the room in your house where you keep aspirin, tylenol and other medicines. Where in that room do you keep them? Are they: (PROBE FOR THE MOST COMMON PLACE)

a. On an open shelf 1
b. In a cabinet 2
c. In a closet 3
d. In a medicine cabinet 4
e. On a night table 5
f. In a drawer 6
g. Or someplace else: SPECIFY _____ . . 7
h. Don't know 8
i. Don't have any 9

In Q6.3 a combination of close-ended and open-ended questions with an unordered response is presented. The use of "or somewhere else" offers the respondent the option of answering an unnamed category. In the event that one response in the "somewhere else" category constitutes 10 percent or more of the overall responses, an additional category should be assigned and post-hoc coding performed. In Q6.3 the number 0 could be used for coding a high frequency response in the "somewhere else" category.

The choice of close-ended versus open-ended format depends primarily upon the type of information elicited, the goal of the research project, and the training and expertise of project interviewers.

Questions designed to elicit attitudes are best written in close-ended format. Attitudes are the most difficult type of information to obtain from a single question and are best adduced from a battery of easily answered items whose responses are viewed collectively to access the attitudes sought.

Belief questions can be written in open-ended or close-ended format depending upon the nature of the question. Questions designed to determine whether individuals hold certain beliefs or to test people's knowledge of facts are often written in open-ended format. The degree to which a belief is held, e.g., strongly agree—strongly disagree, very important—not important at all, is best determined in an ordered close-ended question.

Behavioral questions can be asked in either open- or close-ended format. If open-ended questions are to be used to obtain the number of past occurrences of a particular behavior (e.g., the number of times the respondent has gone to the doctor), researchers should guard against requiring too much precision on the part of respondents unless memory-assistance devices like charts or receipts are used. Precision can be improved by using a time limitation, e.g., last two months, that would place reasonable expectations on the part of the respondent to remember occurrences. Close-ended questions that allow the respondent the choice of a broad range lead to less precision. If continuous data such as the total number of occurrences are to be post-hoc categorized into groups for analysis purposes, it is advisable to use a close-ended question.

Attribute questions are normally written in a close-ended format. The sensitivity of attribute questions can often be reduced through offering categories or ranges as response options.

Question Wording

The final decision in writing a question is considered by many the most precarious: how to word the question. What follows are recommendations that may be useful in overcoming common problems in question writing.

1) Question wording should be simple and clear to all respondents. Avoid vaguely defined concepts such as often, many, and usually that evoke different interpretations in the minds of respondents. Question vagueness is also the product of omitting important modifiers that may reduce ambiguity in the question. In example 7 respondents may not be sure if the question refers to all the children in the household, a majority of the children, or any children at all.

Example 7: Vague Wording

Problem:

Q7.1 Can the children in your household swim well enough on their own to keep their heads above water?

1. Yes 1
2. No 2
3. Don't Know . . . 9

First Revision:

Q7.2 Can some of the children . . . . . .

Second Revision:

Q7.3 Can all of the children . . . . . .

Best Revision:

Q7.4 First, can you tell me the number of children in your household 17 years old or younger?

_____ NUMBER

(If number greater than 0) Next, of the children in the household 17 years old or younger, how many can swim well enough on their own to keep their heads above water?

_____ CHILDREN

Further, ambiguity may arise in respondent's denotation of the word "children". Does this refer to biological offspring of any age or to only those that have not yet reached adulthood? Through splitting the question into two parts and specifying an upper age cutoff, the researcher will be able to calculate the proportion of children in households that are able to swim.

The best revision shown in example 7.4 also incorporates a skip pattern that allows interviewers to avoid asking respondents certain questions based upon responses to previous items. The use of skip patterns is preferable to asking questions that do not apply to a segment of the sample.

2) Avoid the use of double-barreled questions. These questions require respondents to answer more than one question with one response. In example 8 the respondent may be unsure which question to answer. Various combinations such as "nightly operation and Saturdays" or "only Saturdays operation" would receive the same response of "No". The additional time required to answer another question is far outweighed by the unambiguous responses to each question.

Example 8: Double-Barreled Question

Problem:

Q8.1 Does your health department operate clinics at least 1 weeknight per week, but not on Saturday?

1. Yes 1
2. No 2

Revision:

Q8.2 Does your health department operate clinics at least 1 weeknight per week?

1. Yes 1
2. No 2

Q8.3 Does your health department operate clinics on Saturday?

1. Yes 1
2. No 2

---

3) Response categories should be mutually exclusive and exhaustive. Ideally, there should be no overlap in response categories such that answers could fall into more than one category. Consider the response categories in example 9. The highest range of the category is identical to the lowest range of the subsequent category. While most people don't make exactly $10,000 or $15,000, there is a tendency for people to round incomes to the closest $1,000. This rounding could easily place many respondents at the shared income points, probably resulting in an upward bias to income responses.

Example 9: Mutually Exclusive and Exhaustive

Problem:

Q9.1 I am going to read some income ranges. Please tell me which range best represents the **total combined income** of everyone in your household, in 1985, including sources such as wages, salaries, interest, dividends, rented or farm income, and government payments. (READ INCOME RANGE)

a. $2,500 - 4,000 1
b. $4,000 - 6,000 2
c. $6,000 - 8,000 3
d. $8,000 - 12,000 4
e. $12,000 - 16,000 5
f. $16,000 - 20,000 6
g. $20,000 - 30,000 7
h. $30,000 + ... 8
i. Don't Know 9

Revision:

Q9.2 I'm going to read some income ranges. Please tell which range best represents the **total combined income** of everyone in your household, in 1985, including sources such as wages, salaries, interest, dividends, rented or farm income, and government payments. (READ INCOME RANGE)

a. Under $2,500 1
b. $2,500 - 4,000 2
c. $4,001 - 6,000 3
d. $6,001 - 8,000 4
e. $8,001 - 12,000 5
f. $12,001 - 16,000 6
g. $16,001 - 20,000 7
h. $20,001 - 30,000 8
i. Over $30,000 9
j. Don't Know 0

The response categories in the problem question Q9.1 obviously overlap. Corrections could entail using the ranges under 2500, 2500-3999, 4000-5999, etc. The response categories in the problem question are not exhaustive in that the under $2,500 category is omitted. In writing questions that have possible response categories that may not be included in the list provided the interviewer or respondent, the use of a residual "other" category ensures all responses are included.

4) Consider carefully your use of time reference. Behavioral questions are often referenced in time such that the respondent is required to count the number of times, in a specified period, he or she has performed some action. In example 10 the first problem question specifies a period from January 1 to the date of the interview. If the survey covers a two- to three-month period there should be a progression from smaller numbers reported in the earlier interviews and larger numbers in later interviews.

Example 10: Time Referents

Problem:

Q10.1 How many times have you had your blood pressure checked this year?

_____ TIMES

Revision:

Q10.2 How many times in the past 6 months have you had your blood pressure checked?

_____ TIMES

Second Problem:

Q10.3 How many days per week do you play tennis?

_____ DAYS

Revision:

Q10:4 On the average, how many days per week do you play tennis?

_____ DAYS

Second Revision:

Q10.5 During the past week how many days did you play tennis?

_____ DAYS

Another problem often encountered in behavioral questions is the lack of a time reference as depicted in the second problem question.

Some respondents would perceive an implied instruction to average the number of times they played tennis over a year to arrive at a weekly figure. This would, in most cases, probably result in a very small integer if calculated properly. Some respondents would reference the period that they were actively involved in tennis, a warm weather sport, and consequently their responses would be higher than an average over the year. For tennis, an activity with strong seasonal variation, Q10.4 requests the respondent to do some mental arithmetic. Consider the answer a rough approximation. Finally, if a specific time reference is to be used, make sure that the period is long enough to capture some activity. In a survey of the adult population a one-week interval would probably not be sufficient to capture any spelunkers or philatelists.

Questionnaire Construction

The list of preceding suggestions is by no means exhaustive. These problems occur with some regularity in questionnaire design and are, for the most part, easy to fix. After most of the questions have been written, the researcher will have to combine the questions to create the perception of an integrated questionnaire. Consideration must be made for how the survey is to be presented to the respondent.

Questionnaire construction varies by each of three survey types — mail, face-to-face, and telephone. Mail surveys, normally subject to the greatest nonresponse, are constructed to present an interesting and pleasing appearance. A recent questionnaire was mailed to public school teachers and principles in North Carolina who administer the fluoride mouthrinse program. Each questionnaire was printed in a 7 X 8½ inch booklet format with questions reduced to 75 percent of the original size. This allowed several pages of material to be condensed leaving the impression that the questionnaire could be easily completed with a small amount of time. Over ninety percent of the respondents completed the mouthrinse questionnaire in a six-week period.

Special attention in the mail survey is also given to the choice of first question with easy, neutral, and interesting questions stated first. Respondents often form their opinions of a survey on the basis of a quick scanning of the tops of pages, probably stopping on the first question.

Instructions to the mail survey respondent must be explicit, since interviewers are not available for explanation. Enclose instructions in parentheses for each question. This will lead the respondent to check instructions to the question and answer accordingly. Response categories or options available to the respondent in a mail survey can often be more detailed than in the telephone survey because the number of categories in the latter is limited by respondent memory.

The telephone and face-to-face survey questionnaires are similar in construction, both primarily designed for verbal communication. Concern for appearance is replaced by attention to how it sounds (2). Interviewers conducting face-to-face interviews can use visual aids to help explain questions and can perceive misunderstandings through cues from the respondent.

Telephone survey interviewers, in the span of fifteen minutes, may make contact with a potentially hostile respondent, assure the individual of their sincerity, secure his or her cooperation in completing the interview, read, code, and follow skip patterns in the questionnaire, and throughout must maintain a friendly and pleasing disposition. Telephone survey questionnaires must be designed to flow easily from question to question. Hesitation, the result of complicated skipping rules, or awkward question placement, can add time to the interview and increase the risk of a termination. The use of arrows to guide interviewers through difficult skipping patterns, the use of oversized pages to reduce turning, and the color coding of sections of the questionnaire are all useful aids to interviewer performance. Questions should be grouped by subject to enhance the conversational nature of the questionnaire. Each subject group should be prefaced by a transitional statement that alerts the respondent that a change of topic is at hand, e.g., "And finally, I have a few questions that ask for a little more information about your...."

Telephone survey respondents are often called upon to discuss with strangers unfamiliar topics at inopportune times. Anxiety, the result of an unexpected call and, in many cases, a deeply held view that the survey is bogus must be modified quickly to avert an immediate refusal. Just as in the case of the mail survey, the first few minutes of the interview are crucial to respondent compliance. The initial questions should be nonthreatening to all respondents, be easily understood, and pertain to the topic of the survey.

Final Steps

After the agonizing ordeal of wording questions is finished there is one more important task to complete before beginning full scale data collection: the pretest. Question wording should be considered an iterative process of successive approximations often never fully completed until the questionnaire is reluctantly sent to the printer. Initially, pretest the instrument with an independent panel of peers not directly responsible for the questionnaire. Objectivity is the rule at this stage of the design. Following peer review, the questionnaire should be tested by a convenient sample of respondents that match as well as possible the target population. Random selection is not essential, with approximately 10-20 respondents sufficient to reveal poor wording.

Finally be forewarned that as data collection is completed and analysis of the data has begun someone will invariably ask "Why didn't you ask .... ?"

ACKNOWLEDGEMENTS

Helpful suggestions from Dr. William Kalsbeek and the assistance of Jane Stanley and Jean Stafford in preparing the manuscript are much appreciated.

REFERENCES

1. Warwick, Donald P. and Charles A. Lininger, *The Sample Survey: Theory and Practice* (New York: McGraw-Hill, 1975).
2. Dillman, Don A., *Mail and Telephone Surveys The Total Design Method* (New York: John Wiley and Sons, 1978).

STATE LIBRARY OF NORTH CAROLINA
3 3091 00004 2754

STATE LIBRARY OF NORTH CAROLINA
3 3091 00747 2657

Department of Human Resources
Division of Health Services
State Center for Health Statistics
P.O. Box 2091
Raleigh, N.C. 27602-2091
919/733-4728

BULK RATE
U.S. Postage PAID
Raleigh, N.C. 27602-2091
Permit No. 1862

CHERYL W. MCLEAN
ASSISTANT DOCUMENTS LIBRARIAN
DIVISION OF STATE LIBRARY
ARCHIVES/LIBRARY BUILDING
RALEIGH, NC INTEROFFICE

500 copies of this public document were printed at a cost of $75.00 or 15¢ per copy.